University of Montana

ScholarWorks at University of Montana

Graduate Student Theses, Dissertations, & Professional Papers

Graduate School

2005

Evaluation of growth and yield model performance against remeasured permanent plots

Melissa A. Jafvert

The University of Montana

AN EVALUATION OF GROWTH AND YIELD

MODEL PERFORMANCE

AGAINST REMEASURED PERMANENT PLOTS

by

Melissa A. Jafvert

B.S. University of Montana, 2003

presented in partial fulfillment of the requirements for the degree of

Master of Science in Forestry

The University of Montana

May 2005

Approved by:

Kelsey S. Milner

Chairperson

Dean, Graduate School

6-10-05

Date

ABSTRACT

Jafvert, Melissa A., M.S., May 2005 Forestry

An evaluation of growth and yield model performance against remeasured permanent plots.

Chairperson: Kelsey Milner KSM

Managers need estimates of growth and yield model reliability. In an effort to investigate model reliability, this study compares two empirical growth and yield models to an independent dataset, a dataset not used in previous model development or model validation. Model projections for the Forest Projection and Planning System (FPS) and the Forest Vegetation Simulator (FVS) are statistically and graphically compared to remeasured plot data for 87 stands. Statistically, FVS outperformed FPS for predicting trees per acre, basal area per acre, quadratic mean diameter per acre, mean tree height per acre, and cubic foot volume per acre. Model adequacy is defined by invalidation outcomes. Both models performed within defined constraints.

ACKNOWLEDGEMENTS

This project would not have been possible without the support of the Forest Biometrics Research Institute, including Jim Arney and Kelsey Milner. The opportunity to work with them and their never-ending patience was greatly appreciated.

Several other people helped answer my questions and kept me on track throughout the project. I would like thank Renate Bush, from the Region 1 USFS, who provided the data and exceptional support throughout the project, my committee members Peter Kolb and Rudy Gideon for their comments and critiques, Renee Lundberg for helping me with the database extension of FVS, and the FVS support folks in Fort Collins, CO. Lastly, thanks are also due to Chuck Vopicka because he is an excellent office compatriot.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

Introduction

Growth and yield models fill an important niche in the process of planning for future forest resources. Resource managers from both the private and public sectors of forestry rely on these growth and yield models to provide estimates of expected forest growth and yield. With this growing reliance on computer simulators for forest planning, even the role of forest monitoring has shifted to answer not the question of what inventory is on the ground but instead answer whether it was what was expected (Iles 1994).

Currently we are in a time of data explosion because the tools for the evaluation process are greater than ever before (Iles 2003). According to Burk (1987), validation is not only one of least glamorous aspects of modeling but contributes little to acceptance of work in refereed outlets. Coupled with the cost of needing an independent dataset for analyses, validation is justifiably the modeler's last activity. Unfortunately, this aversion results in a general lack of invalidation projects.

Model users need estimates of model reliability for a variety of applications. In an effort to investigate model reliability, this study compares two empirical growth and yield models to an independent dataset, a dataset not used in previous model development or model validation. Though apparently simplistic, this process is the ultimate test for a model (Mills 1987).

The growth and yield models used in the study are the individual tree distance-independent model, the Forest Vegetation Simulator (FVS), and an individual tree distance-dependent model, the Forest Planning and Projection System (FPS). Previous invalidation studies for FPS are internal proprietary company reports not available to the public. Prognosis, the earlier version of FVS, has been subjected to different validation projects, but the work is usually for internal monitoring and is not published (Chad Keyser, Pers. Comm., USDA For. Serv. May 27, 2005).

The purpose of this study is to graphically and statistically compare growth and yield model predictions from the Forest Projection and Planning System (FPS) version 6.44 and the Forest Vegetation Simulator (FVS) version 6.21 against an estimate of truth. Model inputs are raw plot data in the appropriate formats at time 1, and an estimate of truth is the remeasured plot data. Only the control plots are used for the analyses. First it is necessary to understand model growth performance before attempting the complexity of a silviculture analysis.

Both FPS and some FVS variants cover the same geographic range of the data, however structurally the models are very different. FVS uses empirical growth equations for small and large tree growth, but FPS uses a nonparametric strategy for calculating small and large tree growth (Ritchie 1999, Arney et al. 2004). FPS uses site index to represent site productivity while FVS represents site productivity with slope, aspect, elevation, location, and habitat type (Ritchie 1999). Both models have a wide array of users, public and private landowners, throughout southeast Alaska, Washington, Oregon, California, Idaho, Montana, and Hawaii (Ritchie 1999). FVS is the most widely used growth and yield model for managing public lands, whereas a recent survey found FPS to be the preferred growth and yield model for managing private lands (Growth Model Users Group. http://www.growthmodel.org. March 8, 2005) The Northern Idaho (Inland Empire) and Kootenai/Kaniksu/Tally Lake variant of FVS and the current Region 14 of the Species Library in FPS were used in this study

Validation, or as Zuuring et al. (1987) discussed, invalidation of models is dependent on the user, the model, and the dataset. Zuuring et al. (1987) define invalidation of tree growth models as the final step in model building in which one tests the reliability of the model to predict the dependent variable of interest. Brand and Holdaway (1983) agreed, stating validation usually implies the determination of model acceptability while evaluation occurs the entire time the user interacts with the model. Burk (1986) points out that even if models are not completely comparable the observed deviations from actual data can be studied for consistency with expectations. He promotes cautiousness, however, in interpreting validation results with respect to user domain. Newberry and Stage (1987) defined three basic forest growth and yield model user domains: harvest scheduling, inventory updating, and evaluation of silvicultural investments. Each user domain requires different types of data for testing. Long-term data is appropriate for the harvest scheduling user domain, while short-term data is appropriate for testing inventory updating, and data with treatment effects is appropriate for evaluation of silviculture. Results from this study will be interpreted as pertinent to the inventory update user domain.

During the invalidation process there are a few possible outcomes. Newberry and Stage (1987) presented that the validation process ends with one of four outcomes for a particular decision: 1. Model is adequate; 2. Model needs revision using the available data identified in the process; 3. Data appear inadequate to evaluate the model and new data are required; 4. Model is irrelevant. Goulding (1979) states that validation attempts to increase user's confidence in a model, not prove the model is correct. G.E.P. Box supports this stating all models are false but some are useful (Monserud 2003).

Using the validation outcomes presented by Newberry and Stage (1987), model precision and bias will be observed to formulate a statement of model reliability for each growth and yield model. Prediction trends identify model bias. Model precision is user specified and defined by a break-even point such that the decision variable is a difference, presumable zero (Stage and Newberry 1987). The decision point, or break-even point defined by Newberry and Stage (1987) is used to evaluate the adequacy of the model in terms of probability of a wrong decision.

Data

Although there are many models and many datasets, there are few good datasets because good data is difficult to obtain (Curtis and Hyink 1984). Only credible permanent fixed-area plots are appropriate for the purpose of model evaluation. The permanent plots used for this analysis are the Paired Cluster Plots maintained by the USDA Forest Service Region 1. Installed in the early 1980's, the purpose of the paired cluster plots is to look at treatment effects, validate Prognosis (the earlier version of FVS), and to construct yield tables (Bush 2003)

Originally 420 plot clusters were installed, but only 242 are still in the USDA Forest Service Region 1 remeasurement program (Bush 2003). The USDA Forest Service Region 1 maintains the plots since they took over the remeasurement program in 1997. Known problems with the data include control and treated plots being thinned, and some forests deviated from standard Region 1 stand exam protocol (Bush 2003, Region 1 1991). Due to known errors in the data, only 87 control plot clusters are classified as clean and are available for analysis (Renate Bush, Pers. Comm., USDA For. Serv.

December 1, 2004). The 87 plot clusters, with three control plots per cluster, are well distributed among seven national forests (Figure 1).

The plot design includes three control plots and multiple treatment plots in each stand. Only the control plots are used for this study, and each control plot is a 1/20th acre permanent fixed radius plot with three nested 1/100th acre plots. Regeneration was defined as trees with a diameter at breast height less than a specified value, which ranged from 0.2 to 5.0 inches. In addition to the variable definition of regeneration, the definition changed over time. All trees above the specified regeneration diameter at breast height were bored for ages at plot installation. Tree heights were sub-sampled at all remeasurement periods.

As of 2005, there were four remeasurements for each plot with the remeasurement period varying from 14 to 20 years. In some cases, the time between remeasurements was four years, increasing the chance of bias or measurement error of height or diameter growth. Only measurements one and four were used for this project in order to maximize the length of record for FVS (Renate Bush, Pers. Comm., USDA For. Serv. December 1, 2004). FPS growth predictions are unaffected by different growth step increments.

The designated stands are randomly distributed among slope (%), aspect (degrees), elevation (feet), projection period (years), and habitat type (Tables 1 and 2). There is no apparent trend of remeasurement periods between forests. Eight stands in the Bitterroot National Forest, MT, are distributed around 6000’ which represents the highest elevation sample in the dataset. Ten stands in the Panhandle National Forest, ID, are located around 4000’ in elevation, are mostly greater than 30% slope, and have western hemlock, western redcedar, and grand-fir habitat types. The Clearwater National Forest, ID has eight stands which have mostly western redcedar and alpine fir in the overstory. The Flathead National Forest, MT has twelve stands; the stands are mostly less than 20% slope and are scattered between 3500-5500' in elevation. Twenty stands in the Kootenai National Forest, MT are slightly less than or equal to 4500' in elevation. The Lolo National Forest, MT stands are well distributed in terms of slope, elevation, aspect, and habitat type/dominant species on the plots. Finally, the Nez Perce National Forest, ID has ten plots which are mostly less than 30% slope, around 5000' in elevation, and the dominant overstory species is grand fir, although one plot has a western redcedar habitat type.

The dataset is comprised of many young trees. Three quarters of the trees available for analysis, sample size equal to 9,540, have measured height:diameter pairs at time one (Figure 2). In Figure 2, the trees with diameters and no heights were not measured for heights at time one. Only 20% of the individual trees had measured heights greater than 20 ft. at plot installation. Two-thirds of the subset has measured heights greater than 20 ft at time four. Over one quarter of the trees have a 0 inch diameter at time 1 (Figure 3). Slightly less than 10% have a 0 inch diameter at time four. Three quarters of the trees have a diameter around 5 inches or less at both remeasurements. Species composition for the paired cluster plots is primarily lodgepole pine (*Pinus contorta*), Douglas-fir (*Pseudotsuga menziesii*), and western larch (*Larix occidentalis*). Table 3 shows a complete species distribution at time one.

Methods

Once the data is formatted correctly for model input, both growth and yield models require user specifications for appropriate model parameters. FVS uses the keyword system, a set of mnemonic words associated with the data to communicate information used by the model (Dixon 2002). FPS uses a system of setting certain parameters in various Microsoft Access tables in the FPS user interface (Arney et al. 2004).

Forest Vegetation Simulator Setup Procedures

The Region 1 Paired Cluster Plots were prepared in FVS ready format, so no modifications were necessary to run computer simulations in FVS. However, the appropriate model parameters were assigned to each stand simulation with keywords in accordance with USFS guidelines (Renate Bush, Pers. Comm., USDA For. Serv. December 1, 2004). Record tripling helps stabilize random effects and was turned off (Van Dyck 2003). An associated parameter, the diameter growth standard deviation was completely suppressed as well as regeneration (Van Dyck 2003). In order to "grow" the stands to the proper time intervals the first projection cycle was set to 10 years, and the range for the second growth cycle was then defined as four to nine years. The mortality function is a default utility in FVS and was unaffected by user inputs. Using the database extension utility, FVS output moved directly into a database ready for further analysis.

Forest Projection and Planning System Setup Procedures

The Forest Projection and Planning System requires different parameters than FVS for model setup. Aside from transferring all stand information such as slope, elevation, aspect, latitude, longitude, stand acres, etc., the user must specify a site index and a degree of silviculture treatment for each stand.

It was possible to designate site trees in FPS because ages were measured for many trees above breakpoint diameter at time one; oftentimes for trees in the stand but not on the plot. Site trees were selected from the best possible height:age pairs in the dataset. Table 1 displays the range of site indices for the 87 stands. A site tree was designated if and only if the tree had an "open grown" height:age pair near 50 years. For this analysis, each stand had two to four trees across the species range designated as site trees. Although the potential for an upwards bias of site index exists with this methodology, site index field validation was not possible for this analysis. As a result, the site calculations are a major limitation of the project even though they come from the best available data.

FPS requires the user set parameters representing the degree of site preparation, brush control, and animal control. The specified front-end silviculture for the simulations was 0% site preparation, no planting, 0% brush, and 0% pest control. Although there was no site preparation, brush or pest control, these are default specifications for FPS. Clumpiness, a measure of individual tree variance within a plot or variance between plots in a stand, can be specified by the user but for this project was internally calculated in FPS.

Analytic Procedures

Measures of model reliability can be obtained in many ways from different variables. Predicted per acre variables as well as diameter distributions are compared for each stand. Previous analytic procedures are reviewed, and test procedures are described.

Some literature concerning growth model validation suggests using a single test statistic derived from the statistical analysis to gauge model precision or accuracy (Brand and Holdaway 1983, Newberry and Stage 1987, Patterson and Stiff 1987). Measures gauging differences around the mean difference between observed and predicted values are appropriate for some purposes of invalidating models, but biological data is rarely normally distributed about the mean. McQuillan and Sawyer (1983) used paired sample *t* tests to detect model bias for regression coefficients for Lubrecht Experimental Forest Sanders County permanent plots. In *Testing Accuracy*, Freese (1960) suggested the preferred test is the chi-squared test for differences between means. The chi-square goodness of fit test is appropriate for nominal scale data, but goodness of fit for ordered data may often be handled best by the Kolmogorov-Smirnov test (Zar 1999). The forest mensuration measures targeted for this study fall into the ordered data category. The Wilcoxon paired-sample test is a nonparametric analogue to the paired sample *t* test (Zar 1999). Paired-sample significance testing summarizes broad trends in data such as stand summary measures.

The discussion of what tests to do in validation is critical, but so is the discussion of what variables to test during model validation. The Inland Northwest Growth and Yield Cooperative (INGY) provided Patterson and Stiff (1987) with a set of standards for graphical analyses of the data broken into two classifications: whole stand variables and individual tree variables. Zuuring et al (1987) defined a residual as an observed value minus the paired predicted value, and recommended graphing residuals to look for biases or trends in the data, graphing the dependent variable (residual) on the y-axis and the independent variable on the x-axis. Independent variables include: initial average diameter at breast height, initial trees per acre, initial relative density, initial tree top height, and site index (Zuuring et al. 1987). Patterson and Stiff (1987) also recommended projection period as an additional independent variable. According to Zuuring et al. (1987) dependent variables include: ending average diameter at breast height residual, ending top height residual, ending trees per acre residual, and ending stand volume residual as well as individual tree height increment and diameter increment. Stand-level measures are less precise than individual tree measurements. This analysis uses much of the reviewed methodology, but due to the nature of both growth and yield models, analysis modifications are necessary.

For this study, the Wilcoxon paired-sample test coupled with the Kolmogorov-Smirnov test are used to observe general trends of observed and predicted stand-level values. Trees per acre, basal area per acre, quadratic mean diameter, cubic foot volume per acre, and tree top height are targeted for analysis. The Wilcoxon paired-sample test is employed to detect possible bias of observed and predicted per acre attributes. The Kolmogorov-Smirnov test is used to verify per acre value distribution goodness of fit. If data used for analysis is from a randomly selected population, the standard deviation of the residuals can be observed to identify bias (Newberry and Stage 1987). Residual standard deviation is observed as recommended by Newberry and Stage (1987). A residual is defined as observed minus predicted value as recommended by Zuuring et al. (1987).

When the original data at both remeasurement periods is compiled in both models, variables such as trees per acre, quadratic mean diameter per acre, and basal area per acre compile identically. These measures are assumed to be coming from 100% observed trees on the plots, so both models should compile identically. Other stand-level measurements such as calculations of volume and tree top height do not compile identically due to different calculators in each model. In order to compare volume and tree top height, the compiling scheme was standardized. This standardization was done by moving the output individual tree list from FVS at the fourth remeasurement into FPS and compiling the predicted tree records to acquire stand-level variables, eliminating compilation differences for cubic feet per acre, and tree top height. FPS is not yet developed to output individual tree growth. As a result the tree list from FVS was moved into FPS to compile per-acre values and diameter class distributions. FVS grows one tree. FPS grows trees on a spatial basis, and currently the lowest resolution is 0.25 ac. The three plots comprising a plot cluster are each 0.20 acres, so FPS simulates five trees per every tagged tree. Five simulated trees grow differently depending whether the FPS internal stem map generator placed them in a clump or in an opening.

FPS does not output individual trees, but it does output diameter distributions by stand. Rather than compare observed and predicted individual tree pairs, diameter class distributions by stand are compared. The distributions were broken into 11 diameter classes: a 0 inch diameter class for trees with a diameter less than 0.5 inches, a 1 inch class for trees from 0.5 inches to 1.4 inches, a 2 inch class for trees from 1.5 inches to 2.4 inches, etc. up to a 10 inch class. Trees with diameters greater than 9.4 inches were grouped into the 10+ inch class because there were few trees with larger diameters. The predicted and observed diameter classes by stand were compared using the Kolmogorov-Smirnov goodness of fit for continuous data (Zar 1999).

Unfortunately, survivor growth cannot be isolated for the analysis. It may appear preferable to eliminate small trees from the analysis, but the plot design does not track individual tree numbers for small tree growth. As a result, it is impossible to determine what regeneration trees grew or died between time one and four. The situation, complicated by small tree growth, would be worse if regeneration was eliminated.

Graphical displays were used to aid in visualizing and interpreting statistical results. By far, the most popular graphical analyses of growth and yield models is a display of regression analyses for the observed and predicted values (Goulding 1979, Patterson and Stiff 1987, Zuuring and Arney 1985, Zuuring et al. 1987). Biologic data is rarely normally distributed, so nonparametric tests, accounting for non-normal distributions are preferred for analysis of growth and yield model predictions. A cumulative distribution function is a graphical nonparametric alternative to regression analysis (Higgins 2004).

Both traditional scatter plots and cumulative distribution graphs are employed to identify trends in the data. Observed and predicted stand-level values are plotted with cumulative distribution graphs. Cubic foot volume is the most stable per acre variable because it is the sum of individual tree volumes, or in the case of FPS, the mean of the sum of a group of individual tree records. Observed verses predicted cubic foot net change in inventory, defined as cubic foot volume per acre at time one subtracted from cubic foot volume per acre at time four, is plotted. Cubic foot volume residuals (observed minus predicted values) are plotted verses initial stand conditions such as: initial trees per acre, initial basal area per acre, initial quadratic mean diameter, and site index to delineate prediction trends.

Results

Each growth and yield model performed differently for trees per acre, basal area per acre, quadratic mean diameter, tree top height, and cubic foot volume per acre. Table 4 displays results for the Wilcoxon paired-sample test and Kolmogorov-Smirnov test for observed and predicted stand summary measures. FVS outperformed FPS for every per acre value. Some variables such as trees per acre for FPS and quadratic mean diameter for FVS prove that a goodness of fit test can be used in conjunction with a paired-sample test to gain a better understanding of the sample distribution. Table 5 displays summary statistics for observed and predicted trees per acre, quadratic mean diameter, basal area per acre, tree top height and cubic foot volume per acre as well as residual standard error.

Both models under-predict tree top height and over-predict quadratic mean diameter per acre (Figure 4 and 5). Standardizing the compiling scheme, the predicted FVS and FPS top height and the observed tree top height are defined as the mean tree height of the 40 largest (in diameter) trees per acre. Both models performed similarly, showing a bias which under-predicts the observed values. Quadratic mean diameter per acre is another matter. The FPS bias is greater than FVS, but both models over-predicted quadratic mean diameter (Figure 5 and Table 5). This could be due to differences in trees per acre predictions, but in this instance, that is not the case. FVS and FPS under-estimate trees per acre, so trees per acre probably does not account for both models slightly over-predicting quadratic mean diameter (Figure 6).

FPS over-predicts basal area per acre and FVS does not. FPS over-predicts basal area per acre for three quarters of the stands because quadratic mean diameter was over-predicted (Figure 7). The FVS basal area per acre prediction may be balanced by a slight over-prediction of quadratic mean diameter and a slight under-prediction of trees per acre to provide an accurate prediction of basal area per acre.

Statistically, FPS cubic foot volume predictions are different than the observed values for both tests, but FVS predictions are not different for the Kolgomorov-Smirnov test but are statistically different for the Wilcoxon paired-sample test. Figure 8 and Table

5 display the range of the predicted and observed values for cubic foot volumes. FVS consistently slightly under-predicts cubic foot volumes, and there is a definite trend for FPS predictions to over-predict cubic foot volumes. Although the predicted FVS distribution shape was close to the observed distribution, the constant under-prediction was detected by the Wilcoxon paired-sample test (Table 4).

The differences between model predictions and observed cubic foot volume per acre values may also be due to mortality functions as well as natural regeneration occurring on the remeasured plots. Regardless of the model, cubic foot volume net change in inventory, initial cubic foot volume subtracted from predicted or observed cubic feet per acre, was accurate for only a few stands (Figure 9). Both models have checks and balances to limit predictions. Observing the cubic foot volume change over time shows FVS predictions are not as close to observed values as Figure 8 implies. Likewise, FPS predictions are not as far from observed as Figure 8 implies. Points to the right of the line indicate an over-prediction, and points to the left indicate an under-prediction.

Cubic foot volume results are not dependent initial conditions. Cubic foot residuals do not trend across site index (Figure 10). A negative number indicates an over-prediction and a positive number an under-prediction. FPS consistently over-predicts cubic foot volume regardless of the initial trees per acre (Figure 11). Cubic foot volume per acre for both models is variable across the range of quadratic mean diameters (Figure 12). Figure 13 shows no prediction trends between models for cubic foot volume per acre residuals across initial basal area per acre values.

Natural regeneration makes it difficult to only compare stand-level values. Diameter class distributions attempt to highlight whether differences in stand-level values are due to natural regeneration or other factors, but they are not as precise as individual tree comparisons. The Kolmogorov-Smirnov test tests goodness of fit, so it provides only a general measure of the similarity of predicted and observed values. Both models performed well for the diameter class distributions (Table 6). FPS did not predict the diameter class distribution for two stands due to natural regeneration. One stand had over 1,300 new trees per acre due to regeneration between remeasurements, and the other stand had 1,600 new trees per acre due to ingrowth. Both stands had mortality in the larger trees at time four. FVS performs well even for stands with massive ingrowth because FVS retains a proportion of the diameter class distribution in the smaller diameter classes over time. However, one stand had no small trees at time one but had observed regeneration at time four. In this case, FVS over-predicted the diameter classes because it was not able to predict regeneration. Neither model can be held accountable for regeneration effects because the regeneration functions were turned off during model runs to isolate the growth model.

Many stands observed mortality in the larger trees at the fourth remeasurement, but only an individual tree analysis would be able to make definitive statements about large tree growth. It is difficult to determine if this causes FPS to over-predict cubic foot volume growth per acre because FPS over-predicts diameter classes greater than 3" except for the 10" class and FPS under-predicts trees per acre. Based on general model performance, it is likely FVS predicts mortality better than FPS.

Discussion

Many adjustments can be made to both models which affect model performance. FVS keywords were designated under USDA Forest Service Region 1 guidance, and for this project were appropriate (Renate Bush, Pers. Comm., USDA For. Serv. December 1, 2004). FPS inputs were followed under Forest Biometrics Research Institute guidance (Charles Vopicka, Pers. Comm., FBRI, January 26, 2005). Curtis (1994) emphasizes that when several simulation programs of different structure agree on the general nature of trends one can have some confidence in the estimates, but conversely radical differences in growth patterns indicate either weakness in the simulation program or with the data.

Statistical differences do not necessarily translate into practical differences. In order to discuss model reliability, it is first necessary to discuss the possible outcomes of model validation presented by Newberry and Stage (1987). Both models are relevant, so the first question is to evaluate data available for analysis.

The dataset does not test the models across all possible range of forest stand conditions in the inland empire, but instead represents young heavily stocked stands, across the defined range of species, site productivity, slopes, elevations, and aspects. Remember, research plots should not be representative of the entire land base (Iles 2003). The dataset is comprised of many small trees. Ninety percent of the stands have quadratic mean diameters less than 4 inches at time one. Historically, trees less than 6 inches were not even measured on many field installations, resulting in a general lack of data for this size group. The quadratic mean diameter per acre for previous studies such as McQuillan and Sawyer (1983) was greater than 7.5 inches and for Patterson and Stiff (1987) was around 11.3 inches. It is certainly possible, that a different dataset would reveal different results.

The data available for this study is adequate to evaluate each model (Newberry and Stage 1987), but the analytic level of resolution is constrained by both the data and the FPS model. The plot design prohibits the ability to monitor regeneration as it grows onto the larger plot. This is a major disadvantage, because it is not possible to evaluate growth or mortality of regeneration. FPS does not provide individual tree growth data for analysis, so it is not possible to isolate growth and mortality of individual trees. Another limitation of the data is there was no field verification of site trees. An attempt to calibrate predicted FPS top height by adjusting site index occurred after the analysis was complete. When site index was chosen by FPS regression rather than by hand, trees per acre and cubic foot volume was no longer statistically significant. In addition, tree top height was reduced even further, increasing the magnitude of FPS under-prediction. When site index was increased, cubic foot volume per acre and quadratic mean diameter increased. Slight changes in the site index impact stand-level variables due to the complex nature of the model. Site index was chosen from the most appropriate data and was not determined from which number provided the best estimates, so without field verification, the designated site index was most appropriate for the project. Even with limitations imposed by the data, it is possible to evaluate trees per acre, basal area per acre, quadratic mean diameter per acre, and cubic foot volume per acre.

The next possible invalidation outcome determines asks the model needs revision using available data identified in the invalidation process (Newberry and Stage 1987). FVS performs well except it under-predicts tree top height per acre. A previous USDA Forest Service internal validation study for small tree growth recommended including height growth in FVS input files, so FVS can calibrate height for individual trees (Chad

Keyser, Pers. Comm., USDA For. Serv. May 20, 2005). Height growth was not included in input files because it was unavailable at plot installation. As a result, it is indeterminate as to whether the FVS model needs revision, but the results emphasize the possible necessity of including height growth measurements for increased precision of FVS predictions.

FPS results identify an over-prediction of diameter growth and an under-prediction of height growth. The lack of individual tree data makes it difficult to identify the nature of these differences. Model inputs affect model output, so the site index can impact growth results. Sensitivity testing revealed modifying site index, did not change the overall trend of FPS to over-predict diameter growth and under-predict height growth. FPS model structure uses an external Species Library to predict tree growth. FPS may not need revision of its architecture, but results emphasize the possible need for another calibration of the Region 14 Species Library.

Statistical tests are necessary for invalidation testing, but managers also need practical statements about model performance. On average, if two cruisers journey to the same stand in the field, the difference between cruise estimates can be plus or minus 5 to 10 percent (Dr. Kelsey Milner, Pers. Comm., University of Montana, April 8, 2005). The average difference in cubic foot volume per acre per year between observed values and predicted FPS values is -21 cubic feet per acre per year, translating into a 1.5 percent over-prediction per year. The average predicted difference for FVS is ten cubic feet per acre per year, translating into a less than one percent under-prediction per year. Both cubic foot volume predictions are clearly under that which could be found by field verification. Due to young nature of the observed stands, observing trees per acre finds predictions are also within plus or minus 10 to 20 percent. The average difference between the observed values and FPS is 80 trees per acre per year, equivalent to a two percent under-prediction per year. FVS average difference between observed and predicted values is 64 trees per acre, also under-predicting by about two percent per year.

According to Stage and Newberry (1987), the final validation outcome asks whether the model is adequate, and for this study adequacy is in terms of inventory updating. The minimum difference between predicted and observed values is zero, and the maximum acceptable difference is defined by comparing different field measurements. The Forest Vegetation Simulator (FVS) and the Forest Projection and Planning System (FPS) perform within the defined constraints. Model adequacy depends on the decisions for which the model is to be used (Newberry and Stage 1987). Only model users can determine whether these growth models fit their needs.

Figures

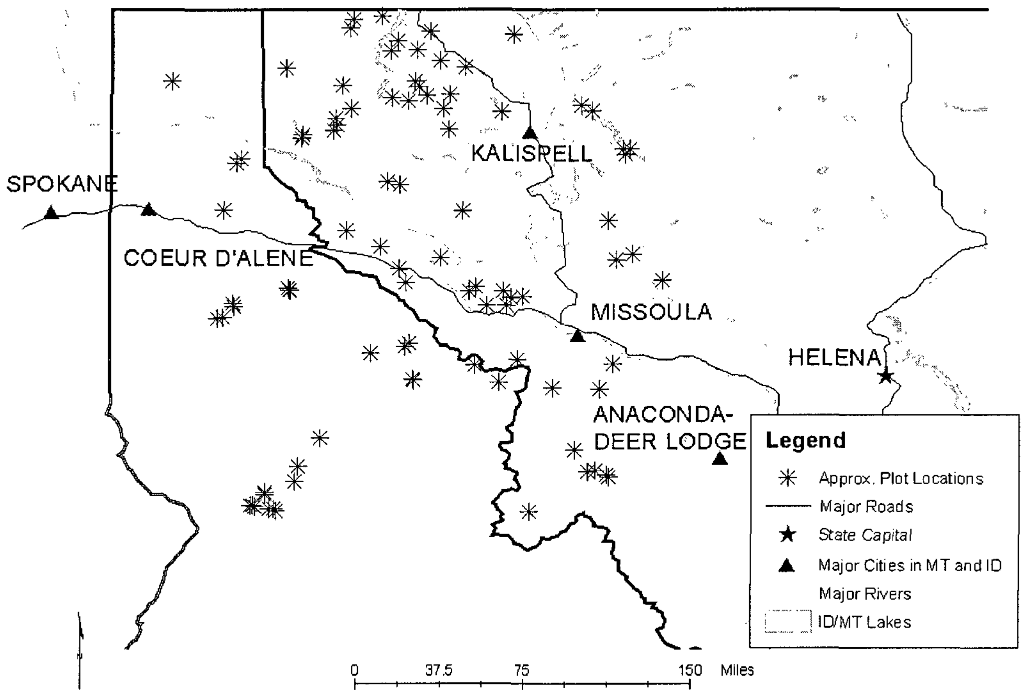

Figure 1. Approximate 87 plot locations throughout western Montana and northern Idaho for the Paired Cluster Plots.

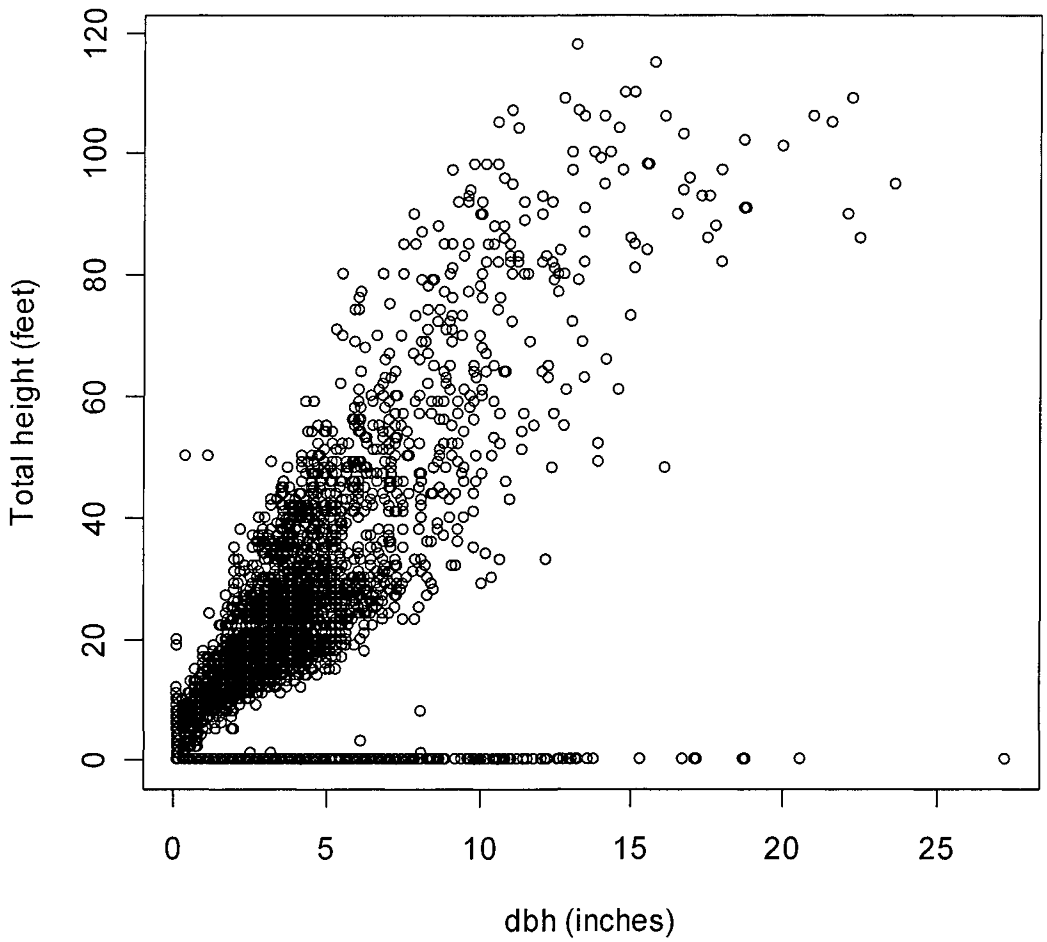

Figure 2. Individual tree height:diameter pairs for 87 stands at time one. Tree heights are subsampled, so zero heights indicate trees not measured for height at time one.

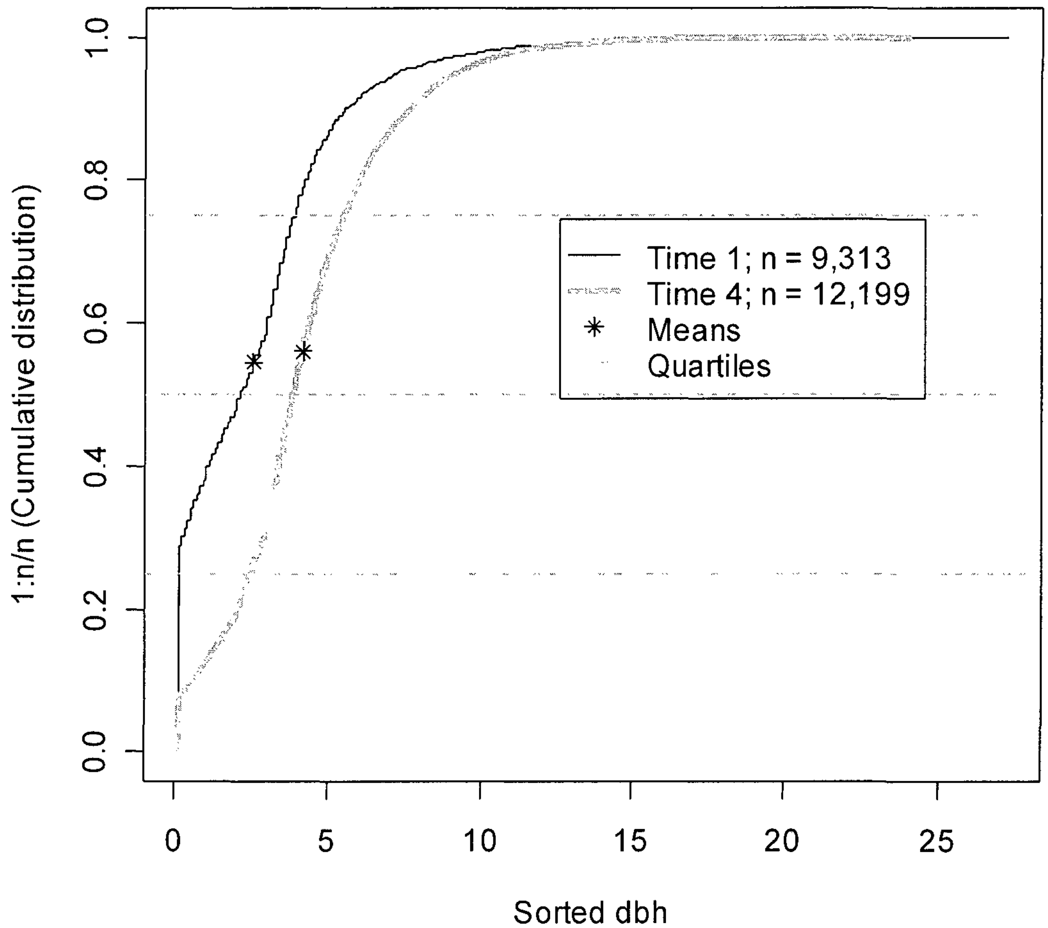

Figure 3. DBH size distributions for Region 1 Paired Cluster Plots. Time 1 is plot installation and Time 4 is the fourth remeasurement of the plots. Over half of the trees are below 5 inches in diameter at both measurement periods. Trees appearing to be zero inches in diameter are actually 0.1 inches in diameter.

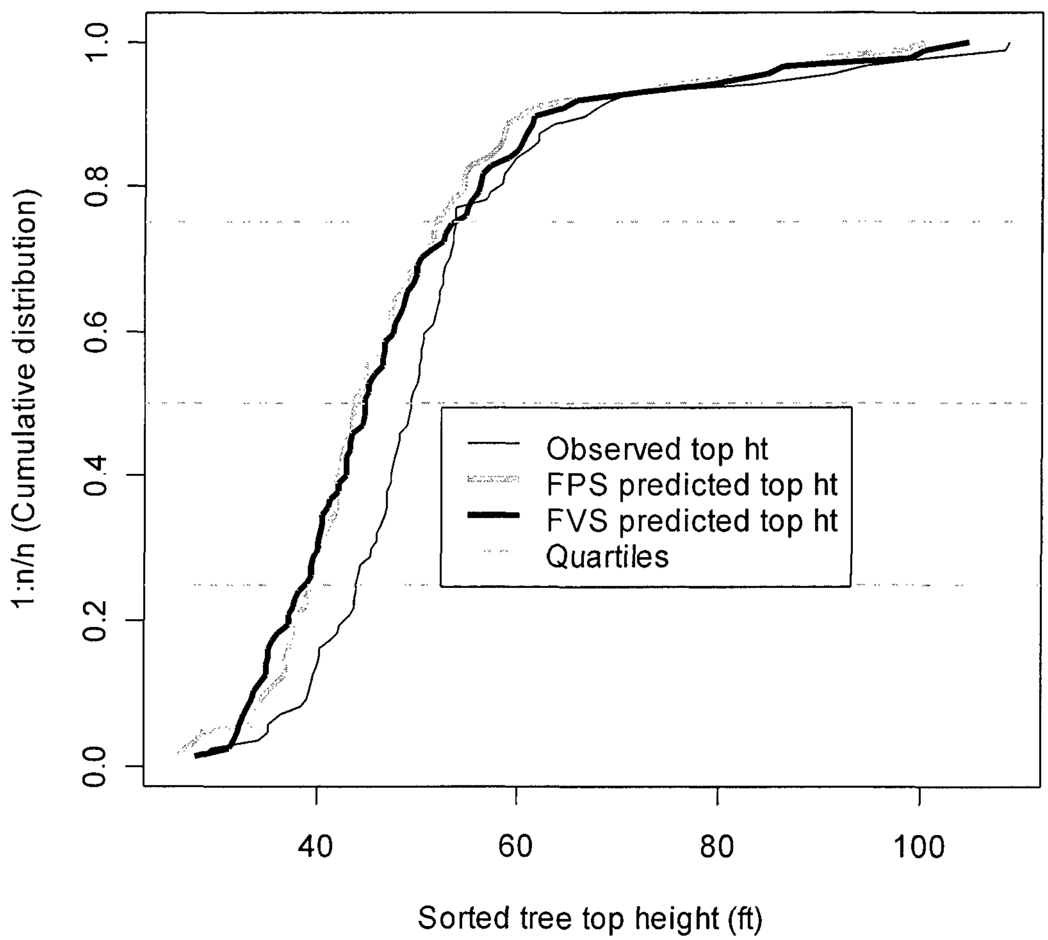

Figure 4. Observed and predicted tree top height.

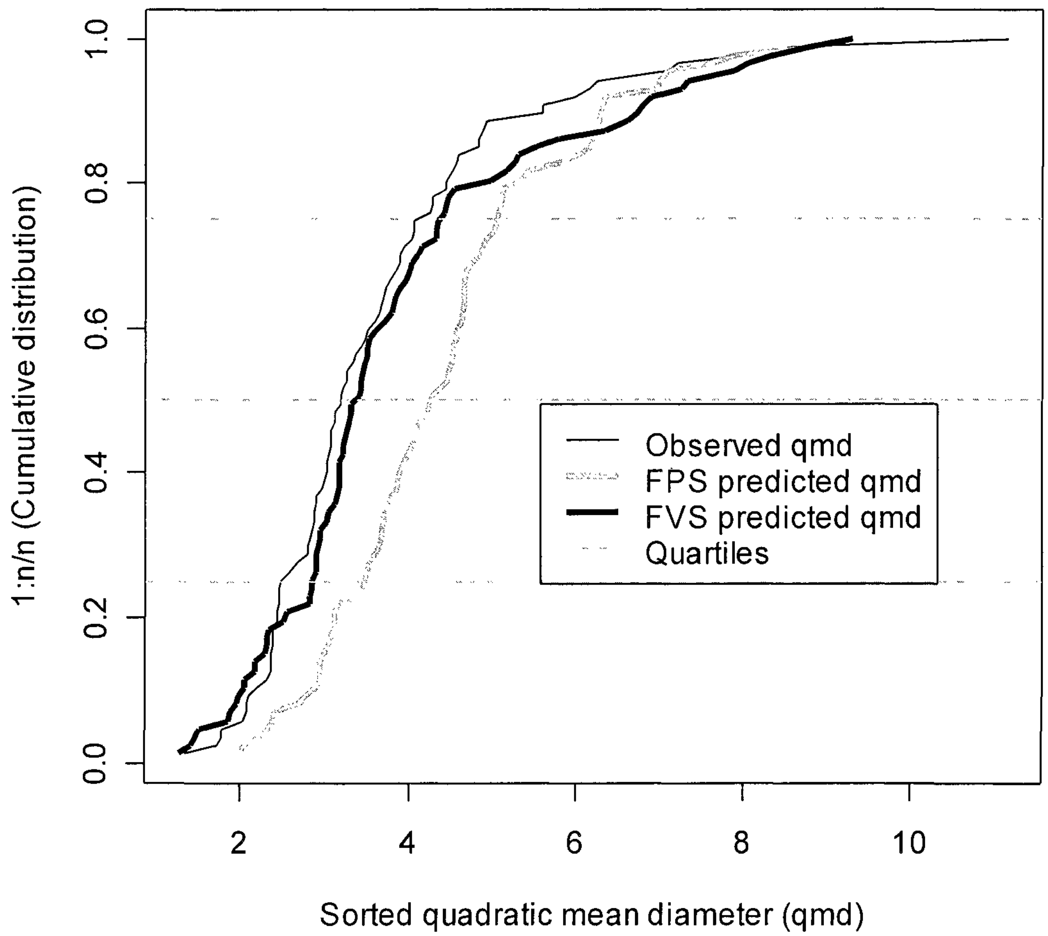

Figure 5. Observed and predicted quadratic mean diameter.

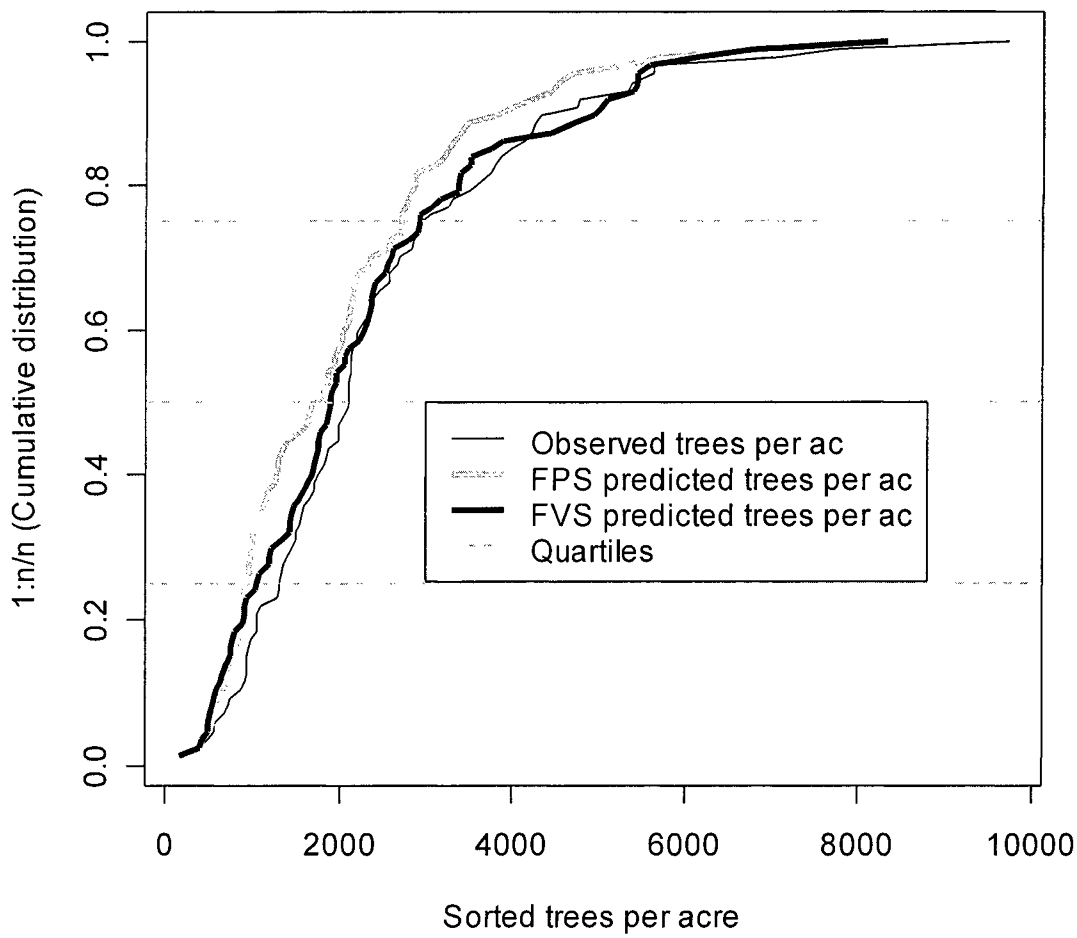

Figure 6. Observed and predicted trees per acre.

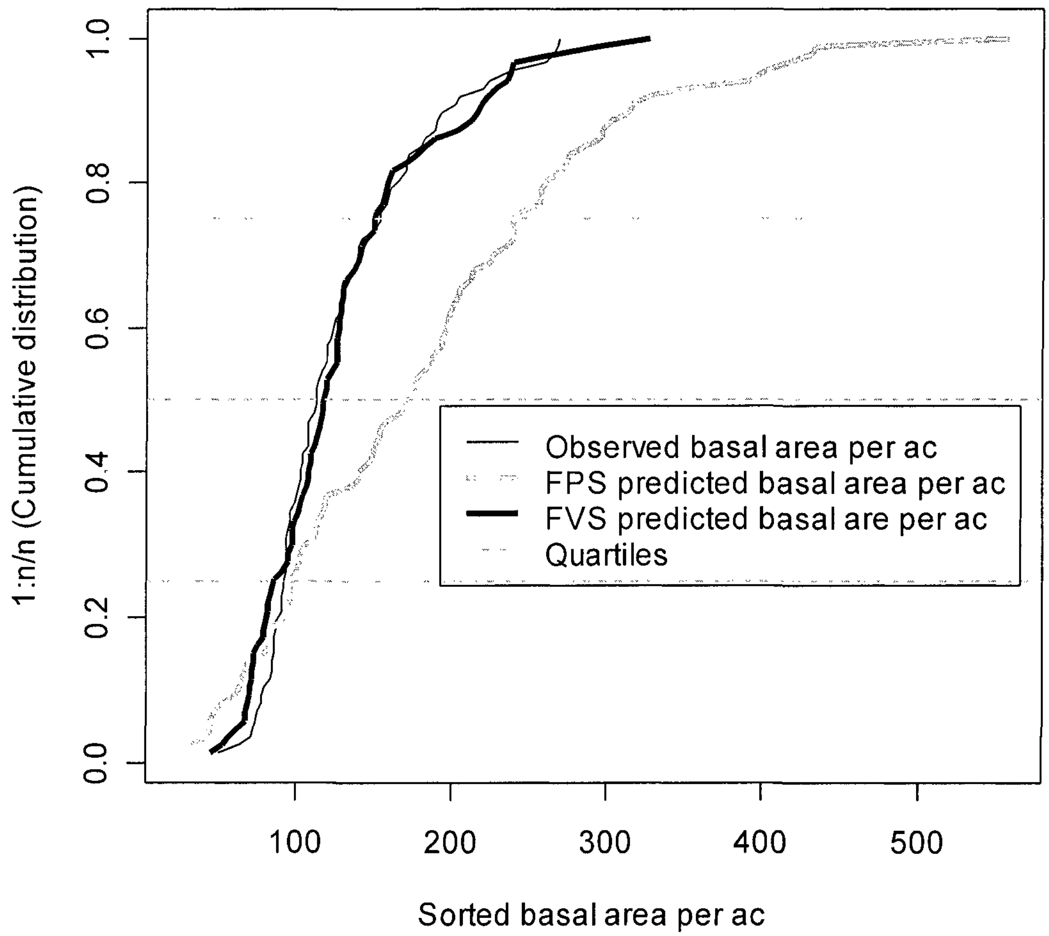

Figure 7. Observed and predicted basal area per acre.

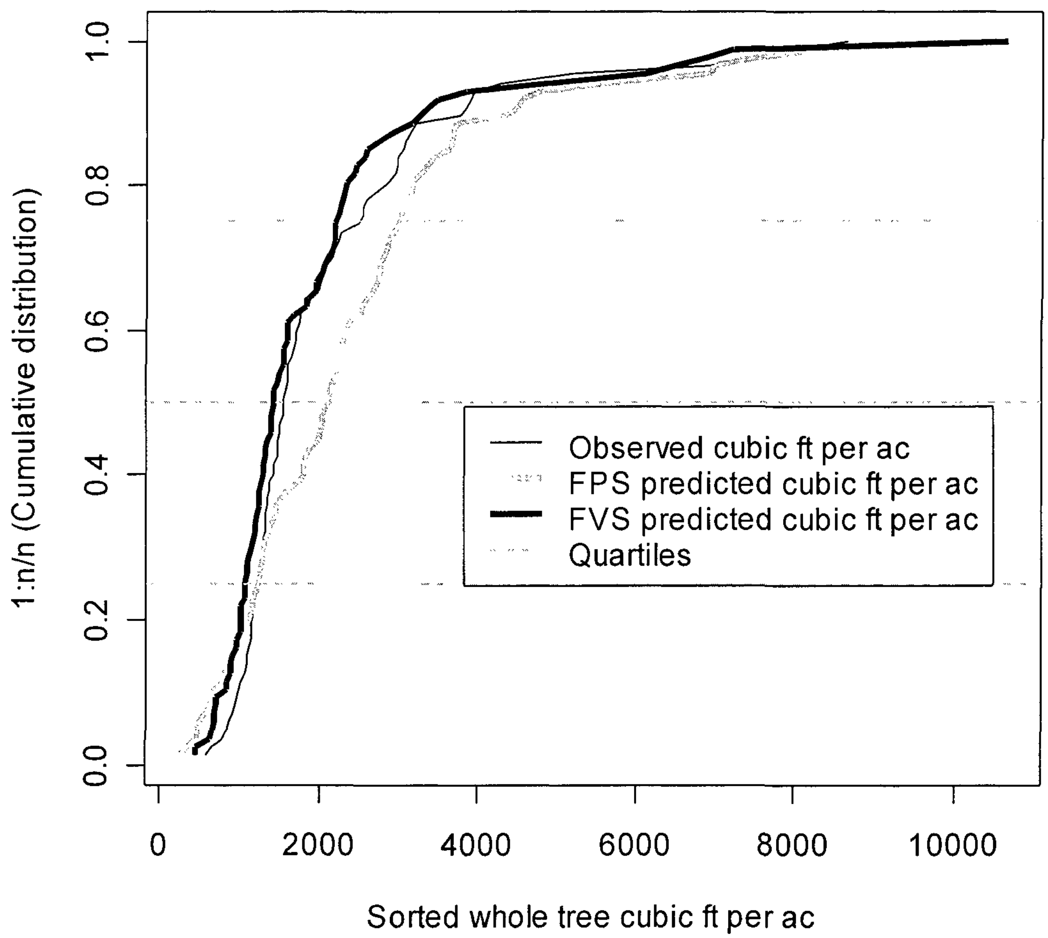

Figure 8. Observed and predicted cubic foot volumes per acre.

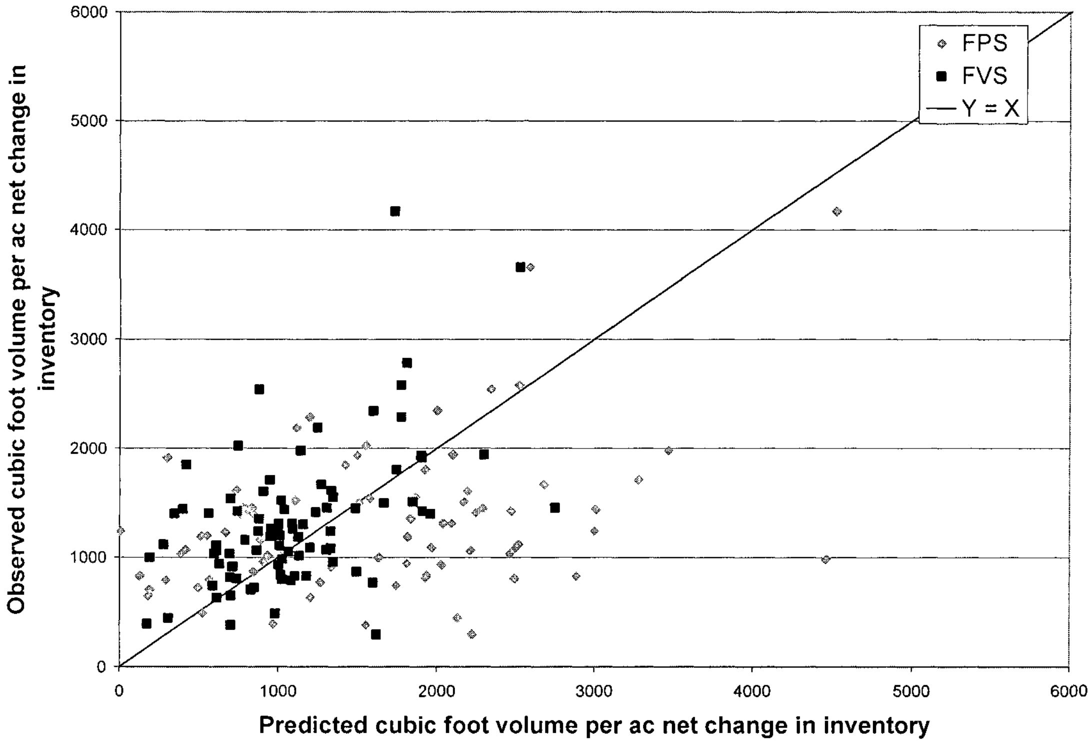

Figure 9. Observed verses predicted net cubic foot volume change in inventory.

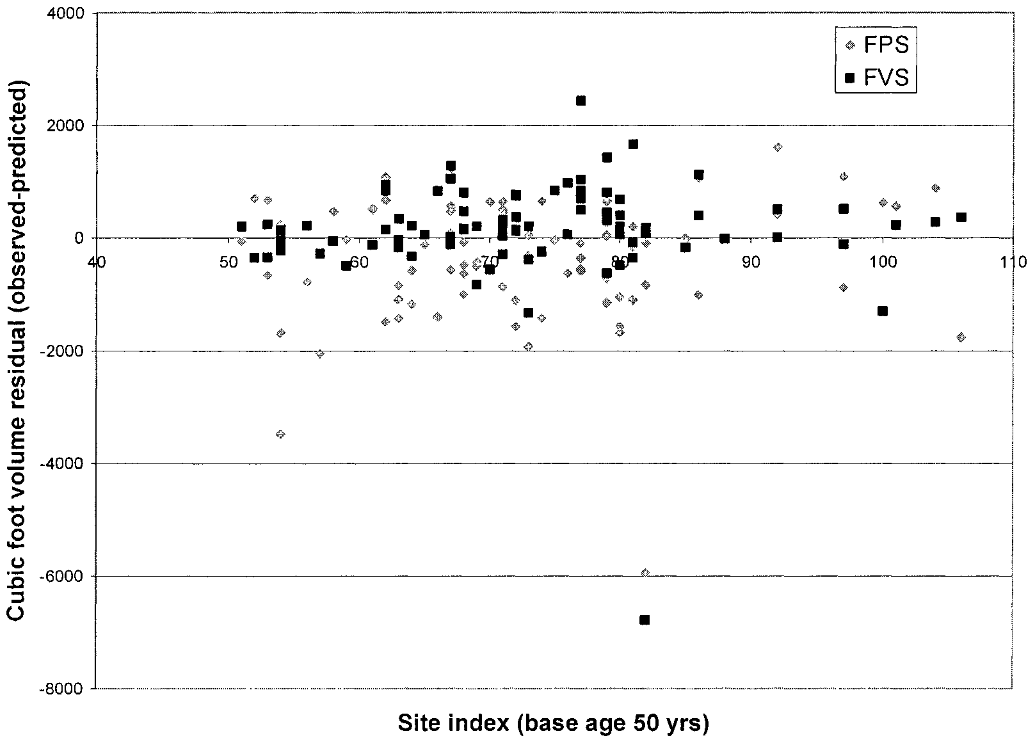

Figure 10. Cubic foot volume residual verses stand site index.

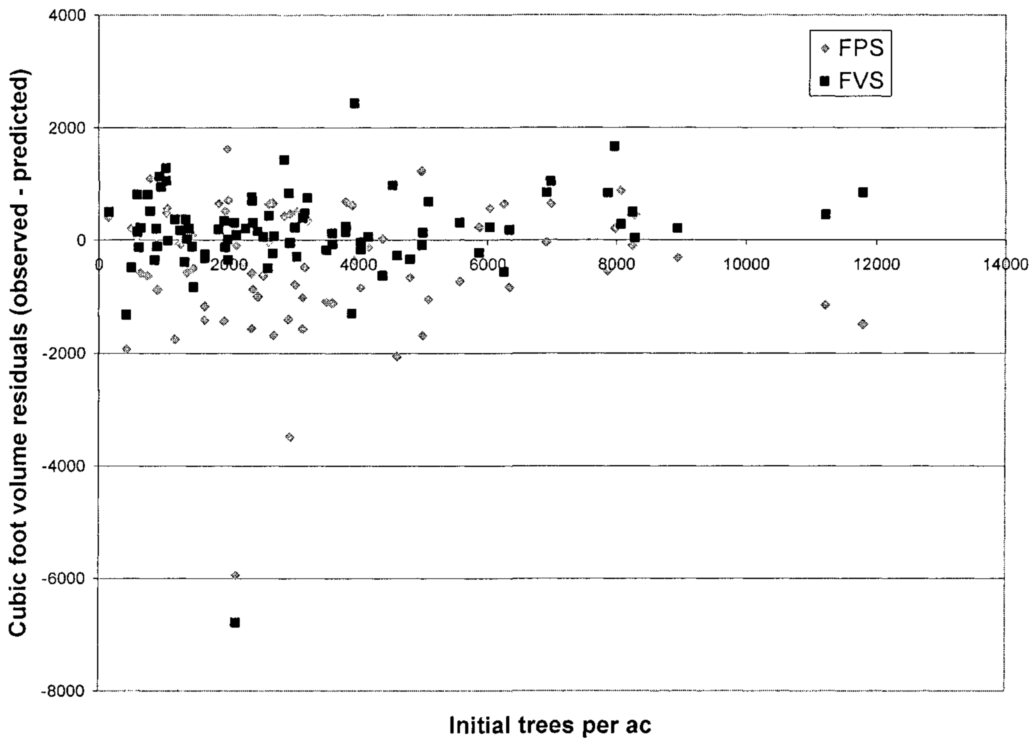

Figure 11. FPS and FVS cubic foot volume residual verses initial trees per acre.

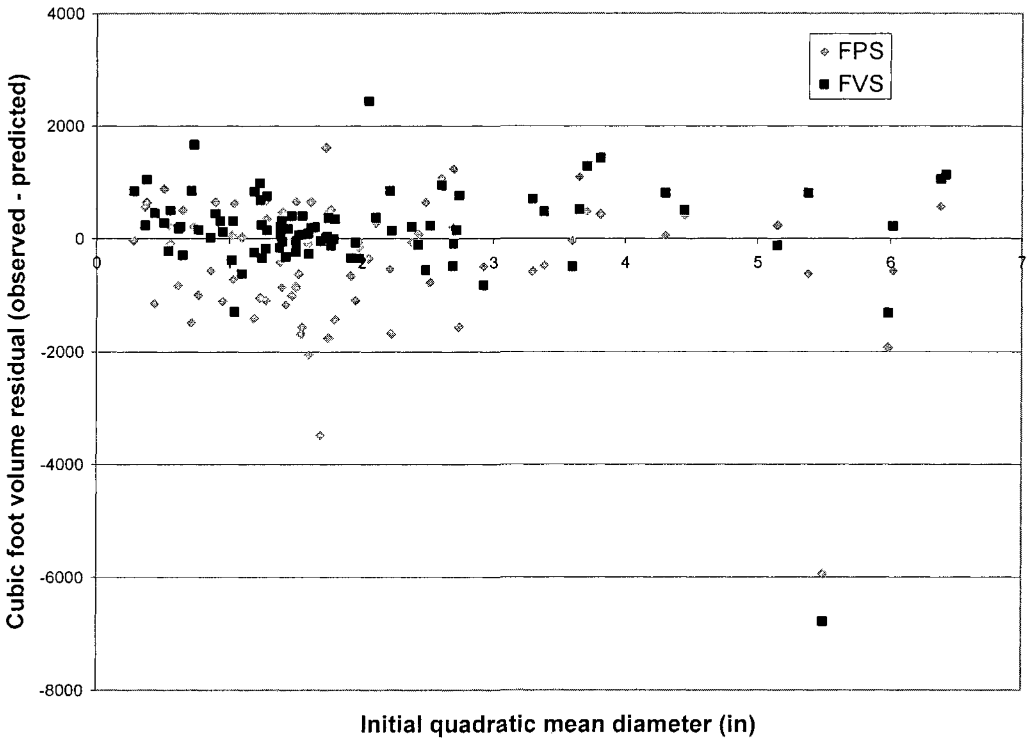

Figure 12. Cubic foot volume residual verses initial quadratic mean diameter.

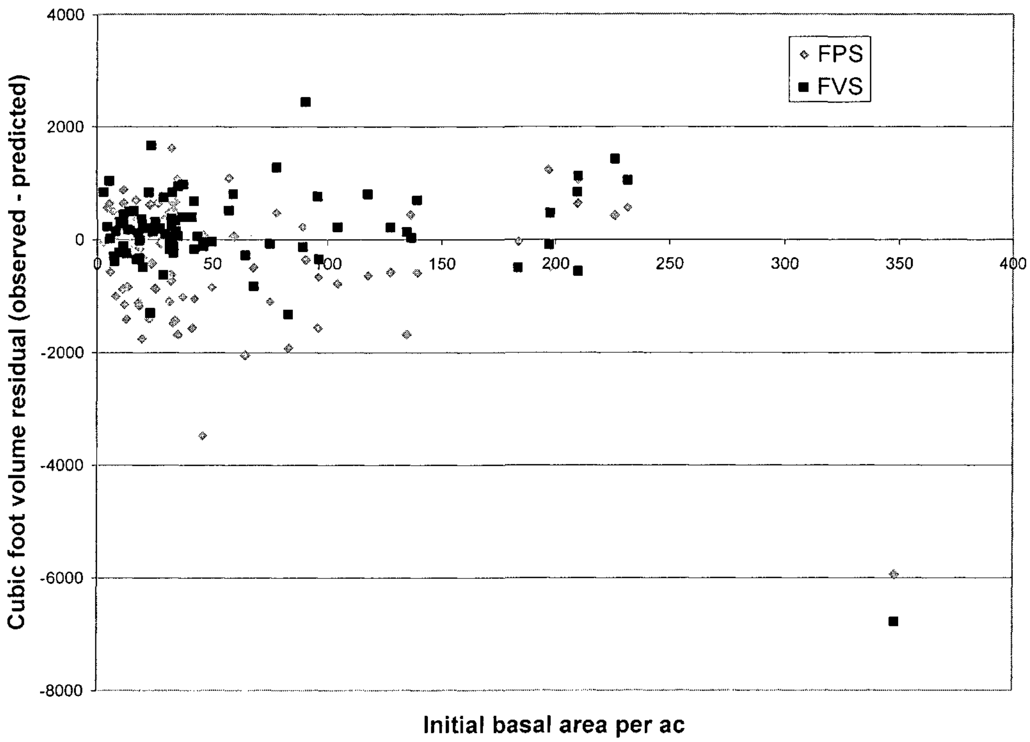

Figure 13. Cubic foot volume residual verses initial basal area per acre.

Tables

Table 1. Stand summary information.

Summary Statistics	Elevation (ft)	Slope (%)	Site Index ($base_{age}$ = 50)	Projection Period (yr)
Minimum	2600	0	51	14
1st Quartile	3700	14.75	65	15
Mean	4507	26.2	73	15.7
Median	4400	24.5	72	16
3rd Quartile	5200	37.25	80	16
Max	6700	63	106	19
Total N:	87	87	87	87
Standard Deviation	933.7	15.7	12.2	1

Table 2. Forest habitat type series frequency in dataset.

Forest Habitat Types Series	Relative Frequency for 87 stands (%)
Lower subalpine series; Abies lasiocarpa (Subalpine fir)	34.5
Abies grandis (Grand fir)	25
Pseudotsuga menziesii (Douglas-fir)	16
Tsuga heterophylla (Western hemlock)	11.5
Thuja Plicata (Western redcedar)	10
Picea spp. (Spruce)	3

Table 3. Paired Cluster Plots individual tree distribution at time one.

No. Stands*	Kootenai NF, MT	Lolo NF, MT	Panhandle NF, ID	Flathead NF, MT	Clearwater NF, ID	Nez Perce NF, ID	Bitterroot NF, MT	Totals	Percent sample by species
	20	19	10	12	8	10	8	87	
Pinus contorta (LP)	239	538	36	576	454	82	363	2288	25%
Pseudotsuga menziesii (DF)	475	326	423	111	128	226	142	1831	20%
Larix occidentalis (WL)	393	185	34	357	6	85	48	1108	12%
Abies grandis (GF)	215	47	336	5	80	209	5	897	10%
Picea engelmannii (ES)	181	170	145	136	87	21	12	752	8%
Pinus ponderosa (PP)	141	195	0	14	0	153	138	641	7%
Abies lasiocarpa (AF)	92	205	45	121	55	0	21	539	6%
Thuja plicata (WRC)	139	66	146	0	68	38	0	457	5%
Pinus monticola (WP)	78	17	111	39	36	0	0	281	3%
Tsuga heterophylla (WH)	69	0	213	0	0	0	0	282	3%
Betula papyrifera (PB)	0	9	0	83	0	0	0	92	1%
Populus spp. (CO)	15	1	0	4	0	11	0	31	<1%
Tsuga mertensiana (MH)	0	23	2	0	7	0	0	32	<1%
Pinus albicaulis (WB)	0	3	0	0	30	0	0	33	<1%
Populus tremuloides (AS)	41	0	0	8	0	0	0	49	<1%
Totals	2078	1785	1491	1454	951	825	729	∑=9313	
Percent sample by National Forest	22%	19%	16%	16%	10%	9%	8%		∑=100%

Table 4. Statistical results for stand-level variables.

	Test pairs	
Per ac variable	Observed values and FPS	Observed values and FVS
Trees per ac	KS $p = 0.072$ Wilcoxon $p = 0.010$ $P(135 \leq d_i \leq 561) = 0.95$ $\Theta = 3337$	KS $p = 0.744$ Wilcoxon $p = 0.614$ $P(-7 \leq d_i \leq 259)$ $\Theta = 119$
Basal area per ac	KS $p < 0.001$ Wilcoxon $p < 0.001$ $P(-66.64 \leq d_i \leq -32.06) = 0.95$ $\Theta = -50.87$	KS $p = 0.492$ Wilcoxon $p = 0.731$ $P(-5.2 \leq d_i \leq 8.00) = 0.95$ $\Theta = 1.07$
Quadratic mean diameter per ac	KS $p < 0.001$ Wilcoxon $p < 0.001$ $P(-1.06 \leq d_i \leq -0.63) = 0.95$ $\Theta = -0.84$	KS $p = 0.381$ Wilcoxon $p = 0.032$ $P(-0.30 \leq d_i \leq -0.02)$ $\Theta = -0.16$
Mean tree height per ac	KS $p = 0.001$ Wilcoxon $p < 0.001$ $P(2.7 \leq d_i \leq 5.5) = 0.95$ $\Theta = 4.1$	KS $p = 0.007$ Wilcoxon $p < 0.001$ $P(2.3 \leq d_i \leq 5.2)$ $\Theta = 3.8$
Crown competition factor	KS $p = 0.001$ Wilcoxon $p < 0.001$ $P(-149 \leq d_i \leq -62) = 0.95$ $\Theta = -107$	$P = 0.859$ Wilcoxon $p = 0.881$ $P(-16 \leq d_i \leq 13) = 0.95$ $\Theta = -1$
Cubic foot volume per ac	KS = 0.01232 Wilcoxon $p = 0.024$ $P(-233 \leq d_i \leq 162) = 0.95$ $\Theta = -29$	KS $p = 0.287$ Wilcoxon $p = 0.001$ $P(80 \leq d_i \leq 319) = 0.95$ $\Theta = 198$

NOTES: All p-values are exact. KS stands for the Kolmogorov-Smirnov test which tests the differences between observed and predicted per ac distributions. Wilcoxon stands for the Wilcoxon paired-sample test for testing the difference between paired observed and predicted per ac values.

Table 5. Summary statistics for observed and predicted stand-level variables.

	Minimum	1st Quartile	Mean	Median	3rd Quartile	Maximum	Standard Error
Observed at time 4							
Trees per ac	208	1310	2437	2087	2974	9734	1695
Quadratic mean diameter	1.33	2.54	3.63	3.19	4.15	11.16	1.61
Basal area per ac	51	93	128	114	154	269	50
Top height	29	44	52	49	54	109	15
Cubic foot volume per ac	597	1226	2110	1564	2517	8661	1547
Predicted FPS							
Trees per ac	146	939	2024	1768	2698	7998	1466
Residual trees per ac							1298
Quadratic mean diameter	1.99	3.48	4.44	4.26	5.05	9.05	1.47
Residual qmd							1.12
Basal area/ac	26	99	184	173	240	557	106
Residual ba/ac							85
Top height	26	40	48	44	52	100	14
Residual top height							7
Cubic foot volume per ac	248	1244	2429	2102	3018	9829	1773
Residual cf/ac							1085
Predicted FVS							
Trees per acre	145	1050	2292	1883	2915	8315	1638
Residual trees per ac							706
Quadratic mean diameter	1.26	2.86	3.86	3.4	4.37	9.3	1.77
Residual qmd							1.06
Basal area per ac	46	89	129	118	151	327	56
Residual ba/ac							40
Top height	28	39	48	45	54	105	15
Residual top height							7
Cubic foot volume per ac	447	1077	1966	1419	2213	10,669	1633
Residual cubic foot volume per ac							950

Table 6. Results from Kolmogorov-Smirnov test for differences between distributions.

H_o: The observed and predicted diameter class distributions are the same.
H_a: The observed and predicted diameter classes distributions are not the same.

	FVS & Observed	FPS & Observed
No. stands Reject H_o:	1	2
No. stands Fail to Reject H_o:	86	85

NOTE: FPS and FVS rejected separate stands.

Literature Cited

Arney, J.D; Milner, K.S, Jafvert, M.A, Vopicka, C.E. 2004. Forest Projection and Planning System Users Guide and Tutorial. FBRI Tech. Rep. 11. 306p.

Brand, Gary J, Holdaway, Margaret R. 1983. Users need performance information to evaluate models. J. For. 81(4):232-34, 254.

Bush, Renate. 2003. Region 1 Permanent Growth Plots USDA FS, Powerpoint presentation.

Burk, Thomas E. 1986. Growth and yield model validation: have you ever met one you liked? P. 35-39 *in* Proc. of Data management issues in forestry: a computer conference and Third Annual Meeting of the Forest Resources Systems Institute. Allen, Sherhonda, Cooney, Timothy M. Forest Resources Systems Institute, Florence, AL.

Curtis, Robert. O. 1994. Some simulation estimates of mean annual increment of Douglas-fir: results, limitations, and implications for management. USDA For. Serv. Res. Pap. PNW-RP-471. 27p.

Curtis, Robert. O; Hyink, David M. 1984. Data for growth and yield models. P. 1-8. *in* Proc of conf. on Growth and yield and other mensurational tricks: a regional technical conference. Logan, UT. USDA For. Serv. Gen. Tech. Rep. INT-193.

Goulding, C.J. 1979. Validation of growth models used in forest management. N. Z. J. of Forestry. Res. 24(1)108-124.

Higgins, James J. 2004. An introduction to Modern Nonparametric Statistics. Thomson Brooks/Cole. Pacific Grove, CA. 366p.

Iles, Kim. 1994. Directions in forest inventory. J. For. 92(12):12-15.

Iles, Kim. 2003. A sampler of inventory topics. Kim Iles and Associates. 869p.

McQuillan, Alan G., Sawyer, Peter S. 1983. Validation of the growth Prognosis model on western Montana permanent growth plots. Review draft on file with FBRI University of Montana. 28p.

Mills, John R. 1987. An evaluation of an inventory-projection system: TRIM model predictions vs. FIA field measurements in North Carolina. P. 770-777 *in* Proc. of IUFRO Forest Growth Modeling and Prediction. Ek, Alan R., Shifley, Stephen R., Burke, Thomas, E. USDA For. Serv., St. Paul, MN.

Monserud, Robert A. 2003. Evaluating forest models in a sustainable forest management context. The FBMIS Group. 1, 35-47.

Newberry, James D., Stage, Albert R. 1987. Validating forest growth models: procedures defined by resource decisions. P 786-793 *in* Proc. of IUFRO Forest Growth Modeling and Prediction. Ek, Alan R., Shifley, Stephen R., Burke, Thomas, E. USDA For. Serv., St. Paul, MN.

Patterson, Donald A, Stiff, Charles T. 1987. An evaluation and comparison of two distance-independent forest projection models in the inland northwest. P. 802-809 *in* Proc. of IUFRO Forest Growth Modeling and Prediction. Ek, Alan R., Shifley, Stephen R., Burke, Thomas, E. USDA For. Serv., St. Paul, MN.

Region 1, USFS. 1991. Field Instructions: Stand Examination Region 1. Timber Management Data Handbook. FSH 2409.21h R1 Chapter 400.

Ritchie, Martin W. 1999. A compendium of forest growth and yield simulators for the Pacific Coast States. USDA For. Serv. Gen. Tech. Rep. PSW-GTR-174. 59p.

Van Dyck, Michael G. 2003. Keyword Reference Guide for the Forest Vegetation Simulator. USDA For. Serv. Forest Management Service Center. 119p.

Zar, Jerrold H. 1999. Biostatistical Analysis. Prentice Hall, Upper Saddle River, New Jersey. 663p.

Zuuring, Hans R, Arney, James D. 1985. Analysis of SPS and PROGNOSIS growth model projections using the Narcisse (WDNR) data set for Douglas-fir. Draft on file at the FBRI library, Univ of Montana. 9p.

Zuuring, Hans R, Arney, James D., Milner, Kelsey S. 1987. Generic graphical analyses for tree growth model invalidation. P 828-834 *in* Proc. of IUFRO Forest Growth Modeling and Prediction. Ek, Alan R., Shifley, Stephen R., Burke, Thomas, E. USDA For. Serv., St. Paul, MN.